

FIG.2A

LOCUS HSU22027 7215 bp DNA PRI 22-OCT-1995
 DEFINITION Human cytochrome P450 (CYP2A6V2) gene, complete cds.
 ACCESSION U22027
 NID g1008461
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
 Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 7215)
 AUTHORS Fernandez-Salguero, P., Hoffman, S.M., Cholerton, S., Mohrenweiser, H.,
 Raunio, H., Rautio, A., Pelkonen, O., Huang, J.D., Evans, W.E.,
 Idle, J.R. et, al.
 TITLE A genetic polymorphism in coumarin 7-hydroxylation: sequence of the
 human CYP2A genes and identification of variant CYP2A6 alleles
 JOURNAL Am. J. Hum. Genet. 57 (3), 651-660 (1995)
 MEDLINE 95397851
 REFERENCE 2 (bases 1 to 7215)
 AUTHORS Fernandez-Salguero, P.
 TITLE Direct Submission
 JOURNAL Submitted (01-MAR-1995) Pedro Fernandez-Salguero, National
 Institutes of Health, 9000 Rockville Pike, Bethesda, MD 20894, USA
 FEATURES Location/Qualifiers
 source 1..7215
 /organism="Homo sapiens"

2/59

FIG.2A CONT.

5'UTR
CDS

782..790
join (791..970, 1237..1399, 2115..2264, 2499..2659,
3207..3383, 4257..4398,4873..5060,5577..5718, 6308..6489)
/gene=CYP2A6V2:
/codon_start=1
/product=cytochrome P450"
/db_xref-PID:g1008462"

/translation=MLASGMLLVALLACLTVMVLMVWQQRKSKGKLPPGPTPLPFIG
NYLQLNTEQMYNSLMKISERYGPVFTIHLGPRRVVVLCGHDAVREALVDOAEESGGRG
EQATFDWVFKGYGVVFSNGERAKQLLRFAIATLRDFGVGKRGIEERIQEESGFLIEAI
RSTHGANIDPTFFLSRTVSNVISSIVFGDRFDYKDKEFLSLLRMMLGIFQFTSTSTGQ
LYEMFSSVMKHLPGPQQQAFQLLQGLEDFIAKKVEHNQRTLDPNSPRDFIDSLIRMQ
EEEKNPNTFYLKNLMMSTLNLFIAGTETVSTTLGYGFLLLMKHPEVEAKVHEEIDRV
IGKNRQPKFEDRAKMPYMEAVIHEIQRFQDVIPMSLARRVKKDTKFRDFFLPKGIEVF
PMLGSVLRDLRFFSNPRDFNPQHFLGEGKQFKRDAFVFPFSIRKRNCFGEGLARMELF
LFFTVMQNFRCLKSSQSPKDDIDVSPKHVGFATIPRNYTMSFLPR

exon

791..970
/gene=CYP2A6V2:
/number=1

exon

1237..1399
/gene=CYP2A6V2:
/number=2

exon

2115..2264

4/59

FIG.2A CONT.

exon	/gene=CYP2A6V2: /number=3 2499..2659 /gene=CYP2A6V2: /number=4 3207..3383 /gene=CYP2A6V2: /number=5 4256..4398 /gene=CYP2A6V2: /number=6 4873..5060 /gene=CYP2A6V2: /number=7 5577..5718 /gene=CYP2A6V2: /number=8 6308..6489 /gene=CYP2A6V2: /number=9 6490..6744	1646 a	2196 c	1746 g	1627 t
3'UTR					
BASE COUNT					
ORIGIN					

5/59

FIG.2A CONT.

BASE COUNT

1	aagttccctt	gaaatatggc	tctggtcttc	ctccccctgc	caatgaagaa	gatggcagtg
61	gaggttctat	ggcagccatc	ctggcctcac	tctgaggttc	caatgaggat	tctgggcac
121	aagagacagc	tctgggcaaa	gctaaatcaa	gtcagccccct	ggacccagtg	ctgggctgct
181	gggtttcttg	ggagaacgcc	gctgggcttg	ctacacactc	ctcctccag	aaactccaca
241	cccacagccc	tgggtcttcc	tagccccgag	acttcaagt	ccatatgcct	ggaatcccc
301	ttcctgagac	ccttaaccct	gcatactcca	caacagaaga	ccccaaatg	cacagccaca
361	ctttgtctta	ccctaataaa	accagacct	ttggattcct	ctccccctga	acccccagat
421	ccgcacaact	ttgggggtgca	ttctcactct	cagaccccaa	atccaaagcc	caagtgcctc
481	cctatgcaaa	tattccaaac	tcctcagttc	tacagcttat	ctgttgcccc	ctcctaaatc
541	cacagccctg	cggcaccctt	cctgaagtac	cacagattta	gtctggaggc	cccctctctg
601	ttcagctgcc	ctgggggtccc	cttatcctcc	cttgctggct	gtgtcccaag	ctaggcagga
661	ttcatggtgg	ggcatgtagt	tgggaggtga	aatgaggtaa	ttatgtaatc	agccaaagtc
721	cacccctctt	tttcaggcag	tataaaggca	aaccacccca	gccgtcacca	tctatcatcc
781	ctctaccacc	atgctggcct	cagggatgct	tctggtggcc	ttgctggcct	gcctgactgt
841	gatgggtctg	atgtctgttt	ggcagcagag	gaagagcaag	gggaagctgc	ctccgggacc
901	caccccatlg	cccttcattg	gaaactacct	gcagctgaac	acagagcaga	tgtacaactc
961	cctcatgaag	gtgtcccaag	acaggagagat	gggtgtctcg	gggtgggggc	tgcctagttg
1021	gctggggctt	tgtggcaggg	ggttgaccag	tgtggaccag	agtcttagga	aatggagttt
1081	tggagtttca	gcatacagaa	gacaggatct	tgggatgtcc	agctccctga	ctgtgagaac
1141	ctgggtgcga	agcatcccag	cacatgacat	ctcgggtgctg	ggccccattc	agagtggagg
1201	gttctccctc	taaccactcc	caccacctc	catcagatca	gtgagcgcta	tggccccgtg
1261	ttcaccattc	acttggggcc	ccggcgggtc	gtggtgctgt	gtggacatga	tgccgtcagg

6/59

FIG.2A CONT.

```

1321 gaggctctgg tggaccaggc tgaggagtc agcgggcgag gcgagcaagc caccttcgac
1381 tgggtcttca aaggctatgg tgcccaagag ggggaagtg ggcaggtgga cacgaaggtc
1441 tcagtgttcc cagccttctc cctgactctc ctgacaactg gaggataagg gagagtcccc
1501 agtctggtct tccctcccca tctccctaca ttggggcctc tccatgtgta tccctcacct
1561 gtctccagcg gccctgtcct gattcctccc tgccctctctc tgccccacct ccttattctc
1621 tctcactgga gtctcctctt tccctctctc tccctctctc ctccatctct tgggtttctg
1681 ttaccagcc ctgggtctct gtctacatga gtcttttgag ccctcttagc ttctgggctt
1741 ctctgggttt ctcatctctc cggatccctt tctcaattct tctctgtct taggatgcca
1801 gggttattcc tacttccaca tcttcaggct ccatctcctg gtaacagtct ctcttcctc
1861 cagacctctc ctgtttctat ctcaatatta aactctctgc tccagctcag cttaagaatc
1921 tcacaccaag agaggatgtc ctccaccag atctcccat atctcactac cccacctcc
1981 atctctgccc tccatcactc tcttctctc ccaactgccc tgcggaagcg atccaatgga
2041 gtgtggagct aatgccgtga agctatgtgc atctctctgt ctggccgtac ctgggtaata
2101 acctgacga ctaggcgtgg tattcagcaa cggggagcgc gccaaagcag agcgcatcca
2161 tgccatcgcc acctgaggg acttcggggt ggcaagcga cggagcacg aggggacccc
2221 ggaggagtcg ggcttcttca tcgaggccat cggagcacg aggcgtgag aggggacccc
2281 gagtgcgggg gcaggagaag gaaaacaccc aggacgagga acccgcgcg gtctgcctg
2341 gggatgggga ctaggtgagg aaaggcgccc gcaactccag ccctggagtc tggcgctggg
2401 aatttggctc aacaaggccc tgctctctgg aatctgact ctctcagac ctctgagttg
2461 actctctccc caacccccct ctcccgacat acccgaggc gccaatatcg atccacctt
2521 ctctctgagc cgcacagtct ccaatgtcat cagctccatt gtctttgggg accgctttga
2581 ctataaggac aaagagttcc tgtcactgtt gcgcatgatg ctaggaaatct tccagttcac
2641 gtcaacctcc acgggggcagg taatggttgc agccggccc gtgaaggccc ttacaaaac

```

7/59

FIG.2A CONT.

2701 cggcaaatg ttccctacc ggggaaggg ggcccaaat tccaccgcc cccggacag
 2761 tgtccctca aatcagtc cggatttggg caaattggca gagtgaacc agaccgggt
 2821 tggttgtcca atccctgct ctccaggac accgggatag cacaacagat gctcccaaa
 2881 acagagcctg ctggcaggat gcataccctc agctcagctc tctaccctg ggcacgtgtt
 2941 cccatccca acttaccgtt aatttctaac agatgctccc taccaggtc tcttgaata
 3001 ttttaacac cggaaccct gggtacctaa ccttcctgt aaactttaga gattagtcc
 3061 tatccggccc ctctgaaata cctaaccacc ggagaccaga tgcctttaac tcagttcctt
 3121 ccttgctatg aaacaaatcc cattcccatc agtccttgcc ccgtgacagc tgccttccc
 3181 ttcccatcct ctcttgcaa cccagctct atgagatgtt ctctcggtg atgaaacacc
 3241 tgccaggacc gcagcaacag gccttctcagt tctgcaagg gctggaggac ttcatagcca
 3301 agaaggtgga gcacaaccag cgcacgctgg atcccaatc ccacgggac ttcattgact
 3361 cctttctcat ccgcatgcag gaggtagacc ccagcagcca ctgcggggag atgcaaaagcc
 3241 aggcagaggg aaatcagctt gggagtgggg caggcagatg acacaggccc attcaaatla
 3481 accctcatca taataatcct cacaattggc tgggtgccgt ggctaacagc ctgtaatccc
 3541 agcactttgg gaggccgagg caggtggatc acctgaggtc aggagttcga gaccagcctg
 3601 gccaacatgg tcaaaccccg tcttactaa aaatccaaa attagttggg catgggtggcg
 3661 cgaagggggg cagaggttgc aatgagccaa gatcacggca ttgcaactcca gtctgggtga
 3721 cagaatgagg ccctgtgtca aaaaaatta atcacttggt taaaaagtaa gtgagcctgc
 3781 atggtcatgc gcatgtgcag ctccagctac tcaggaggct gaggtggag gattgcttga
 3841 gctcaggagt tggcgtccgg cctgtgcaac ttagcaagac caagtcagta taagaaaaaa
 3901 aaaaaacaaa aaaaagctg acagctaaatgataaattga cggacagatg gtcagcaagg
 3961 taacgaaggt gagaagggaag agcattgggg gcaacgccag ggtcagggc aagggtggt

8/59

FIG.2A CONT.

4021 tcctagagcg agtctggtag gatctagggc ccctcttctc caccctgcgg tcttgcccca
 4081 aagagaggtc gaggtgctg ggattgcgt agactcgagt ctgtgtagat ctgggggtcc
 4141 cctcttgacc ccattggtc tgaacctaa agtggaaagt ccatgggggtg aacccttaga
 4201 tggtgccctg aggtcaagca ggagtgggt tgtcctaaag cccctctcc cttcaggagg
 4261 agaagaaccc caacacggag ttctacttga agaacctgat gatgagcac ttgaacctct
 4321 tcattgcagg caccgagac gtcagacca ccctgcacta tggcttctta ctgctcatga
 4381 agcaccacga ggtggagggt aaggctggag ggggacggaa gtggagggcc ccagaccctc
 4441 aaaattcccc ttgactggt gcaatgtccc cactgtccc agatccccgg accctgagac
 4501 gtgacttgct gtccagagac agggcaacat tcagctggta ggcatacagt gagtctcat
 4561 agatatataa atattgaaa tgctgcact gattggtcag tcaactctgt ccaagccca
 4621 ctgagtgcgc actgcccgtt ccaccgggtc atcccctaa tccctccctg tgcctcccc
 4681 gtgatttctg cacaacctgg ttaacaggat cctactccaa caatgcgaat gggtgatgtc
 4741 tgttctgtta tgaatgctct acttccgtct cataggcga ggcatttcat ccaccccat
 4801 ttgcctatcc ggactatcat ttctgctct gagaccccta gatacctaaa cacattcccc
 4861 ctccctcccc agccaaggct catgaggaga ttgacagagt gatcggaag aaccggcagc
 4921 ccaagtctga ggaccgggcc aagatgccct acatggaggc agtgatccac gagatccaaa
 4981 gatttgagga cgtgatcccc atgagtctgg ccgcagagt caaaaaggac accaagtctc
 5041 gggatttctt cctccctaa gtgctatccg ccccaacccc ccagactacg gggactccag
 5101 cccctctctg tgtccccagc atcccccca cattagaagc ttcttagacc ctgtcccat
 5161 ccctcaatca gtcaaaaaag acttccccaa ccaccacatc cgttccacct ttccacttag
 5221 aactcctga gtcctgcac tctccagact ctttgttca ggagaatcaa acacatgttc
 5281 ccaaaacttc tatcttaaga aacagaagcc cctttccat tcggcctttt gtcataggga
 5341 cagaaatctc aggtccccc aactcctgcc tagaaggaca tggaccccat gtctcccaa

9/59

FIG.2A CONT.

5401 cttcctgttt cagagatgtg aaccttctat cccccaaggt cctccctcag aggtcccaaa
 5461 ttcccatgcc tgccacttcc cctcacccggg gcaccctagt tccccctcca gcccctgtgt
 5521 actctcaaca atcccccaac ccgcctcatc acatacacct tcctcctccc tcccaggcca
 5581 tagaagtgtt ccctatgttg ggctccgtgc tgagagacct caggttcttc tccaaccccc
 5641 gggacttcaa tccccagcac ttccctgggtg agaaggggca gtttaagaag cgtgatgctt
 5701 ttgtgccctt ctccatcagt aagagaccac tgtttgggtgc caggcttact actcacacca
 5761 gcaggggcct cccttaccga gttccccctct ctgccgtgta gcctagtatt tcccagctt
 5821 ggcaagtcc tgtagcaat ctaccgtcga gccaccaggt gatactccct taactaccaa
 5881 gcaccagta cctgtgccc ggcataaagg aaggaacacat cataccccct tcagaggcgg
 5941 gggaaaacca aagggcagag agaatacagag atttatttcc ctagggtcac acaggagatt
 6001 cttcagcat cctaaaaagg agatgacggc acagcaggtc atatttggga gttcttatct
 6061 gggggaagg ggatcttaaa cctcccattg tggacacctg gcatacgatca acccatctt
 6121 ttggtcatct ttgggtcac tcaaggaaac tgagggtcaag gaggtcaag aggtccctc
 6181 ttaaagtctc tcaggggcat atattccacc ctctctccct gggagagccg cagctggagg
 6241 tcggtactgg ggcgaggctg cactgagagt gggcttcacc tccacccctc ccgctctcc
 6301 tcctcaggaa agcggaactg ttctggagaa ggcctggcca gaatggagct ctttctcttc
 6361 ttcaccaccg tcatgcagaa ctctccgctc aagtcctccc agtcacctaa ggacattgac
 6421 gtgtcccca aacacgtggg ctttgccacg atcccacgaa actacaccat gagcttcttg
 6481 ccccgctgag cgagggtgtt gccggtgaag gtctgggtggg cggggccagg gaaagggcag
 6541 ggccaagacc gggcttggga gaggggcgca gctaagactg ggggcaggat ggcggaaaagg
 6601 aaggggcgtg gtggctagag ggaagagaag aaacagaagc ggctcagttc accttgataa
 6661 ggtgcttccg agctgggatg agaggaaagg aaccttaca ttatgctatg aagagtagta

10/59

FIG.2A CONT.

6721 ataatagcag ctcttatttc ctgagcacgt acccccggtg cacctttgtt caaaaacccat
 6781 tgcacgctca cctaatttgc cacaacaccc ccttcgaagg ggcgttcattg ccatttttac
 6841 acgtgacaaa actgaggctt agaaagtgtg ctctgatgtc tcacaaaaca taagtgcaca
 6901 gaaaatctgc gaacacagat ctgtgcccac agccttctag acagattctt aaaaagcacc
 6961 tattcctcac gaaaacacgt ttagtataga atcacatggc ctgaacatcc ctgtccgggg
 7021 gagttcccca gagacctggg ggggtggttg cctgccttca ctgcacacat gccacactc
 7081 tcacctactc aacatgctgt gactacccgg gtgtaatctg tgcttgctac cagataaggc
 7141 cactgtagcc cattcagagt cagcccaggg acacaacgag acatgactgg acatacaggg
 7201 tcagtccatt aacaa (SEQ ID NO:1)

11/59

FIG.2B

LOCUS HSP452B6 1415 bp RNA 29-MAY-1992
 DEFINITION Human MRNA FOR CYTOCHROME P-450IIB6.
 ACCESSION X13494
 NID g35206
 KEYWORDS Cytochrome; cytochrome P450IIB6.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
 Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1415)
 MILES, J.S.
 REFERENCE Direct Submission
 AUTHORS Submitted (10-NOV-1988) Miles J.S., Imperial Cancer Research Fund,
 TITLE Lab of Molecular Pharmacology and Drug Metabolism, Hugh Robson
 JOURNAL Building, George Square, Edinburgh, EH8 9XD
 REFERENCE 2 (bases 1 to 1415)
 AUTHORS Miles, J.S., McLaren, A.Q. and Wolf, C.R.
 TITLE Alternative splicing in the human cytochrome P450IIB6 gene
 generates a high level of aberrant messages
 JOURNAL Nucleic Acids Res. 17 (20), 8241-8255 (1989)
 MEDLINE 90045947
 COMMENT The sequence is a compilation of genomic and cDNA clones. **map:
 chromosomal location=19q12-13.2;
 Data kindly reviewed (13-NOV-1989) by Miles, J.S.
 FEATURES Location/Qualifiers

12/59

FIG.2B CONT.

```

source      1..1415
             /organism="Homo sapiens"
misc-feature 9..110
             /note=exon 1, partial"
misc-feature 111..273
             /note=exon 2"
misc-feature 274..423
             /note=exon 3"
misc-feature 424..584
             /note=exon 4"
misc-feature 585..761
             /note=exon 5"
misc-feature 762..903
             /note=exon 6"
misc-feature 904..1091
             /note=exon 7"
misc-feature 1092..1233
             /note=exon 8"
misc-feature 1234..1415
             /note=exon 9", coding region"
BASE COUNT   341 a   430 c   328 g   316 t
ORIGIN
1 gaattccgcc ctgcaccat gaccgctcc caccaggcc ccgcccctg ccccttttgg

```

13/59

FIG.2B CONT.

61 gaaaccttct gcagatggat agaagaggcc tactcaaatc ctttctgagg ttccgagaga
 121 aatatgggga cgtcttcacg gtacacctgg gaccaggcc cgtggtcatg ctgtgtggag
 181 tagaggccat acgggaggcc cttgtggaca aggtgagggc cttctctggc cggggaaaaa
 241 tcgccatggt cgaccattc ttccggggat atggtgtgat ctttgccaat ggaaccgct
 301 ggaagggtgt tcggcgattc tctgtgacca ctatgaggga cttcgggatg gaaagcggga
 361 gtgtggagga gcggattcag gaggaggctc agtgtctgat agaggagctt cggaaatcca
 421 agggggccct catggacccc accttcctct tccagtccat taccgccaac atcatctgct
 481 ccatacgtctt tggaaaacga ttccactacc aagatcaaga gtccctgaag atgctgaact
 541 tgttctacca gactttttca ctcatcagct ctgtattcgg ccagctgttt gagctcttct
 601 ctggcttctt gaaatacttt cctggggcac acaggcaagt ttacaaaaac ctgcaggaaa
 661 tcaatgctta cattggccac agtgtggaga agcaccgtga aacctggac ccagcgccc
 721 ccaaggacct catcgacacc tacctgctcc acatggaaaa agagaaatcc aacgcacaca
 781 gtgaattcag ccaccagaac ctcaacctca acacgtctc gctcttctt gctggcactg
 841 agaccaccag caccactctc cgctacggct tcctgctcat gctcaatacctcatgttg
 901 cagagagagt ctacaggag attgaaacagg tgattggccc acatcgccct ccagagcttc
 961 atgaccgagc caaatgcca tacacagagg cagtcattcta tgagattcag agattttccg
 1021 accttctccc catgggtgtg cccacattg tcaccaaca caccagcttc cgagggtaca
 1081 tcatccccc aa ggacacagaa gtattttctca tcttgagcac tgctctccat gaccacact
 1141 actttgaaa accagacgcc ttcaatcctg accactttct ggatgccaat ggggcactga
 1201 aaaagactga agcttttata ccttctctt tagggaagcg gatttgtctt ggtgaaggca
 1261 tcgcccagc ggaattgttc ctcttctca ccaccatct ccagaactc tccatggcca
 1321 gcccctggc ccagaagac atcgatctga cccccagga gtgtggtgtg ggcaaatatc
 1381 cccaacata ccagatccgc ttccctgccc gctga (see ID NO: 2)